

RESULT 4
AAU98024
ID AAU98024 standard; peptide; 26 AA.
XX
AC AAU98024;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human readthrough acetylcholinesterase unique region.
XX
KW Human: acetylcholinesterase; single-chain variable fragment; scFv;
KW AChE-S; heavy chain variable region; muscle re-innervation;
KW progressive neuromuscular disorder; muscle re-innervation; readthrough;
KW myasthenia gravis; neuromuscular junction abnormality;
KW Eaton-Lambert disease; muscular dystrophy; amyotrophic lateral sclerosis;
KW ALS; post-traumatic stress disorder; PTSD; multiple sclerosis; Dyscrania;
KW post-stroke sclerosis; post-injury muscle damage;
KW excessive re-innervation.
XX
OS Homo sapiens.
XX
PN WO200246422-A1.
XX
PD 13-JUN-2002.
XX
PF 22-MAY-2001; 2001WO-IL000464.
XX
PR 04-DEC-2000; 2000IL-00140071.
XX
PS (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI Soreq H, Flores CF, Nissim A;
XX
DR WPI: 2002-463832/49.
XX
PT Nucleic acid sequence coding for a single-chain variable fragment (scFv)
PT antibody that has specific affinity for the synaptic variant of
PT acetylcholinesterase (AChE-S), useful for diagnosing a neuromuscular
PT disorder, e.g. Myasthenia gravis.
XX
PS Disclosure; Fig 1; 73pp; English.
XX
CC The invention relates to a nucleic acid sequence coding for a single-
CC chain variable fragment (scFv) antibody that has specific affinity for
CC the synaptic variant of acetylcholinesterase (AChE-S), where the scFv
CC antibody consists essentially of a polypeptide comprising the binding
CC portion of the heavy chain variable region of an antibody. Also included
CC are an expression vehicle comprising a nucleic acid sequence coding for a
CC scFv antibody that has specific affinity for the synaptic variant of AChE
CC -S, an scFv antibody specifically recognising and binding to the synaptic
CC variant of AChE-S and a method for the diagnosis of a progressive
CC neuromuscular disorder in a mammal, comprising obtaining a sample from
CC the mammal and detecting intensified expression of at least one of the
CC AChE variants in the sample. The single-chain Fv antibody is useful for
CC diagnosing a progressive neuromuscular disorder which involves any one of
CC muscle distortion, muscle re-innervation and neuromuscular junction (NMJ)
CC abnormalities. The disorder is Myasthenia gravis (preferred), Eaton-
CC Lambert disease, muscular dystrophy, amyotrophic lateral sclerosis (ALS),
CC post-traumatic stress disorder (PTSD), multiple sclerosis, Dyscrania, post-
CC stroke sclerosis, post-injury muscle damage, excessive re-innervation,
CC or post-exposure to AChE inhibitors. The present sequence represents the
CC unique region of human acetylcholinesterase encoded by a readthrough
CC splice variant
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 147; DB 5; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GMOGPAGSGWEGSGSPGVTPLFSP 26

DB 1 GMOGPAGSGWEGSGSPGVTPLFSP 26
RESULT 5
AAW74588
ID AAW74588 standard; protein; 31 AA.
XX
AC AAW74588;
XX
DT 21-DEC-1998 (first entry)
XX
DE Amino acid sequence of the human AChE variant 7.
XX
KW Nuclease resistance; inhibition; human; acetyl-cholinesterase; AChE;
KW central nervous system; CNS.
XX
OS Homo sapiens.
XX
PN WO9839486-A1.
XX
PD 11-SEP-1998.
XX
PF 06-MAR-1998; 98WO-US004503.
XX
PR 06-MAR-1997; 97US-0040203P.
XX
PS (YISS) YISSUM RES & DEV CO.
XX
PI (KOHN/) KOHN K I.
XX
DR Soreq H, Seidman S, Shohami E;
XX
WPI: 1998-506377/43.
XX
PT Treatment of injury to central nervous system - by administration of
PT inhibitor of acetyl-cholinesterase production.
XX
PS Disclosure; Page 62; 88pp; English.
XX
CC This is the amino acid sequence of a human acetyl-cholinesterase (AChE)
CC variant used in the method of the invention, where inhibitors of AChE are
CC used to treat injury to the central nervous system (CNS). The AChE
CC inhibitor can also be used to facilitate transplantation of neuronal
CC cells to the CNS of a patient. The inhibitor can also be used to improve
CC hippocampal neuron survival following injury to the CNS. The CNS injuries
CC that can be treated with the method include epilepsy, stroke,
CC Huntington's disease, head injury, spinal injury, pain, Parkinson's
CC disease, myelin deficiencies, neuromuscular disorders, neurological pain,
CC amyotrophic lateral sclerosis, Alzheimer's disease, and affective
CC disorders of the brain
XX
SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GMOGPAGSGWEGSGSPGVTPLFSP 26
DB 6 GMOGPAGSGWEGSGSPGVTPLFSP 31
RESULT 6
AAW68146
ID AAW68146 standard; protein; 31 AA.
XX
AC AAW68146;
XX
DT 05-OCT-1998 (first entry)
XX
DE Human AChE splice variant E1-4-14-E5.
XX
KW Nuclease resistant; acetylcholinesterase; human; myasthenia gravis; AChE;

KW	Parkinson's disease; Alzheimer's disease; central nervous system;
KM	neuromuscular junction; cholinergic signalling; brain.
OS	Homo sapiens.
PN	MO9826062-A2.
XX	
PD	18-JUN-1998.
XX	
PF	12-DEC-1997; 97WO-US023598.
XX	
PR	12-DEC-1996; 96US-003526P.
PR	13-FEB-1997; 97US-003777P.
RR	02-MAY-1997; 97US-00850347.
XX	
PR	21-JUL-1997; 97US-0053334P.
PA	(YISS) YISSUM RES & DEV CO.
PA	(KOHN/) KOHN K I.
PI	Soreq H, Seidman S, Eckstein F, Friedman A, Kaufer D,
DR	WPI, 1998-348522/30.
XX	
PT	Synthetic nuclease resistant antisense oligodeoxynucleotides - directed
PT	against acetylcholinesterase, useful for treating Parkinson's and
FT	Alzheimer's diseases and myasthenia gravis.
XX	
PS	Disclosure; Fig 12; 89pp; English.
XX	
CC	This represents the amino acid sequence of a human acetylcholinesterase
CC	(AChE) splice variant. The invention provides sequences shown in AAIV41278
CC	to AAIV41285 that represent synthetic nuclease resistant antisense
CC	oligodeoxynucleotides which are capable of selectively modulating human
CC	acetylcholinesterase (AChE) production. These oligonucleotides are
CC	targeted to a splice junction in a splice variant of AChE mRNA and are
CC	capable of selectively modulating human AChE production in the central
CC	nervous system and neuromuscular junction. The invention also provides a
CC	method for determining the efficacy of these human AChEs specific
CC	antisense oligonucleotides. These antisense oligonucleotides can be used
CC	to restore balanced cholinergic signalling in the brain, particularly
CC	related to learning and memory as well as stress disorders, Parkinson's
CC	and Alzheimer's disease. They can also be used to reduce production and
CC	therefore deposition of AChE in the neuromuscular junctions of patients
CC	with e.g. myasthenia gravis. The oligonucleotides work effectively at low
CC	doses while avoiding many of the side effects associated with tacrine and
CC	related cholinergic drugs for Alzheimer's disease and pyridostigmine and
CC	related drugs for myasthenia gravis
XX	
SQ	Sequence 31 AA:
	Query Match 100.0%; Score 147; DB 2; Length 31;
	Best Local Similarity 100.0%; Pred. No. 1.5e-10;
	Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1 GMGGPAGSGMEGSGSPGVTPLFSP 26
DB	6 GMGGPAGSGMEGSGSPGVTPLFSP 31
ID	AAB50035 standard; peptide; 37 AA.
AC	AAB50035;
DE	14-MAR-2001 (first entry)
XX	Acetylcholinesterase readthrough peptide ARP-2.
XX	ARP-2; haemostatic; acetylcholinesterase; AChE; cell growth; human;
KM	cell differentiation; thrombocytopenia; post-irradiation condition;
XX	post-chemotherapy condition; blood loss; stress-induced male infertility.

```

OS Homo sapiens.
XX WO200073427-A2.
XX 07-DEC-2000.
XX 31-MAY-2000; 2000WO-IL000311.
XX 31-MAY-1999; 99IL-00130224.
XX 02-SEP-1999; 99IL-00131707.
XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX Soreq H, Eldor A, Deutch V, Grissau D;
XX WPI; 2001-061523/07.
XX New regulatory peptides having cell growth and cell differentiation
XX activity derived from the C-terminal region of acetylcholinesterase
XX useful in promoting growth, survival and differentiation of stem cells.
XX Example 10; Page 76; 133pp; English.
XX
XX The present invention relates to C-terminal peptides of
XX acetylcholinesterase (AChE) (see AAB50032-B50034). The peptides of the
XX present invention have cell growth and/or cell differentiation activity.
XX The peptides may be used in ex vivo or in vivo expansion of
XX hematopoietic stem cells and neural progenitors, and in the promotion of
XX megakaryocytic differentiation of hematopoietic stem cells. In addition,
XX the peptides may be used in for promoting expansion of committed neural
XX progenitors in a developing embryo, in cultured embryonic stem cells, and
XX embryoid bodies derived from them. The peptides may further be used in
XX the treatment of thrombocytopenia, post-irradiation conditions, post-
XX chemotherapy conditions, and conditions following massive blood loss, in
XX inducing synthesis of AChs mRNA, and in promoting formation of hematoma
XX bodies. Antibodies directed against the peptides are useful for
XX diagnosing stress-induced male infertility. The present sequence is a C-
XX terminal AChE "readthrough" peptide (ARP-2)
XX
XX Sequence 37 AA;
SQ
Query Match 100.0%; Score 147; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GMOGPAGSGMEEGSGSPGCVPLPFLSP 26
| | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12 GMOGPAGSGMEEGSGSPGCVPLPFLSP 37
RESULT 8
AAB50036
ID AAB50036 standard; protein; 53 AA.
XX AAB50036;
XX
XX 14-MAR-2001 (first entry)
XX
XX Acetylcholinesterase protein #1 used in a yeast two-hybrid system.
XX
XX ARP; haemostatic; acetylcholinesterase; AChE; cell growth; human;
XX cell differentiation; thrombocytopenia; post-irradiation condition;
XX post-chemotherapy condition; blood loss; stress-induced male infertility.
XX
XX Homo sapiens.
XX
XX WO200073427-A2.
XX
XX 07-DEC-2000.
XX
XX 31-MAY-2000; 2000WO-IL000311.
XX
XX 31-MAY-1999; 99IL-00130224.
XX

```

```

PR 02-SEP-1999; 99IL-00131707.
XX
XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Soreq H, Eldor A, Deutch V, Girsaru D;
XX
XX WPI; 2001-061523/07.
XX
XX New regulatory peptides having cell growth and cell differentiation
XX PT activity derived from the C-terminal region of acetylcholinesterase
XX FT useful in promoting growth, survival and differentiation of stem cells.
XX
XX
XX Claim 8; Page 87; 133pp; English.
XX
XX The present invention relates to C-terminal peptides of
XX CC acetylcholinesterase (AChE) (see AAB50032-B50034). The peptides of the
XX CC present invention have cell growth and/or cell differentiation activity.
XX CC The peptides may be used in ex vivo or in vivo expansion of
XX CC haematopoietic stem cells and neural progenitors, and in the promotion of
XX CC megakaryocytic differentiation of hematopoietic stem cells. In addition,
XX CC the peptides may be used in for promoting expansion of committed neural
XX CC progenitors in a developing embryo, in cultured embryonic stem cells, and
XX CC embryoid bodies derived from them. The peptides may further be used in
XX CC the treatment of thrombocytopenia, post-irradiation conditions, post-
XX CC chemotherapy conditions, and conditions following massive blood loss, in
XX CC inducing synthesis of AChE mRNA, and in promoting formation of hematon
XX CC bodies. Antibodies directed against the peptides are useful for
XX CC diagnosing stress-induced male infertility. The present sequence is a C-
XX CC terminal AChE "readthrough" protein (ARP), which was used in a yeast two-
XX CC hybrid system, to screen for ARP binding partners
XX
XX Sequence 53 AA;
XX
XX Query Match 100.0%; Score 147; DB 4; Length 53;
XX Best Local Similarity 100.0%; Pred. No. 2.7e-10;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GMOGPAGSGWBERGSGSPGVTPLPSP 26
XX Db 28 GMOGPAGSGWBERGSGSPGVTPLPSP 53
XX
XX RESULT 9
XX ABB31331
XX ID ABB31331 standard; protein; 53 AA.
XX
XX ABB31331;
XX
XX 05-NOV-2002 (first entry)
XX
XX GFP-fused AChE variant expression construct, pGARP related protein.
XX
XX Nervous system; drug assay; acetylcholinesterase; AChE; brain;
XX KW Isoform variance; AChE blocker; muscarinic receptor; M1; M2;
XX KW pyridostigmine; muscarinic receptor blocker; scopolamine;
XX KW M1 receptor blocker; pirenzepine; anxiety; post-traumatic stress;
XX KW Alzheimer's disease; muscle malfunctioning; neurodegenerative disorder;
XX KW xenobiotic damage; panic; neuromuscular disorder; Parkinson's disease;
XX KW Huntington's chorea; muscle fatigue; multiple chemical sensitivity;
XX KW autism; multiple sclerosis; Sjogren's disease; GFP; pGARP;
XX KW green fluorescent protein.
XX
XX Unidentified.
XX
XX WO200240994-A2.
XX
XX 23-MAY-2002.
XX
XX 14-NOV-2001; 2001WO-IL001051.
XX
XX 14-NOV-2000; 2000US-0247970P.
XX
XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX

```

```

XX
XX Soreq H, Meshorer E, Sklan E, Shoham S;
XX
XX WPI; 2002-490152/52.
XX
XX Evaluating effect of drugs on nervous system by comparing effect of drug
XX PT on acetylcholinesterase, AChE activity in brain of test animal following
XX FT challenge by AChE blocker and comparing it with control group.
XX
XX
XX Example; Page 52; 114pp; English.
XX
XX The present invention relates to a method and system for evaluating an
XX CC effect on the nervous system of a test drug. The method comprises
XX CC comparing the effect of the drug on acetylcholinesterase (AChE) catalytic
XX CC activity or isoform variance in a brain of a test animal following a
XX CC challenge by an AChE blocker or a blocker of AChE and muscarinic
XX CC receptors M1 and M2 (e.g. pyridostigmine) and comparing this effect with
XX CC that of a known agent, preferably a non-selective muscarinic receptor
XX CC blocker (e.g. scopolamine) or a specific M1 receptor blocker (e.g.
XX CC pirenzepine). The method is useful for evaluating an effect on the
XX CC nervous system of a test drug, including drugs for the treatment of
XX CC anxiety conditions, post-traumatic stress, Alzheimer's disease, muscle
XX CC malfunctioning, neurodegenerative disorders, damage resulting from
XX CC exposure to xenobiotics, panic, neuromuscular disorders, Parkinson's
XX CC disease, Huntington's chorea, muscle fatigue, multiple chemical
XX CC sensitivity, autism, multiple sclerosis and Sjogren's disease. The
XX CC present sequence represents a protein described in relation to green
XX CC fluorescent protein (GFP)-fused AChE variant expression construct pGARP
XX CC in the examples of the present invention
XX
XX Sequence 53 AA;
XX
XX Query Match 100.0%; Score 147; DB 5; Length 53;
XX Best Local Similarity 100.0%; Pred. No. 2.7e-10;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GMOGPAGSGWBERGSGSPGVTPLPSP 26
XX Db 28 GMOGPAGSGWBERGSGSPGVTPLPSP 53
XX
XX RESULT 10
XX AAM48797
XX ID AAM48797 standard; protein; 600 AA.
XX
XX AAM48797;
XX
XX 07-OCT-1998 (first entry)
XX
XX Human acetylcholine esterase-14 readthrough splice variant.
XX
XX Human acetylcholine esterase-14 readthrough splice variant; AChE-14; CNS;
XX KW blood/brain barrier; BBB; 14 peptide; antibiotic; brain tumour; glioma;
XX KW chemotherapeutic drug; central nervous system.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Region 1..574
XX FT /note="This region is encoded by exons 1-4 of AChE"
XX FT 575..599
XX FT /note="14 peptide encoded by intron 4 of AChE; this
XX FT sequence is claimed by the inventors under claim 2 in the
XX FT specification"
XX FT 600
XX FT /note="residue encoded by exon 5 of AChE"
XX
XX WO9822132-A2.
XX
XX 28-MAY-1998.
XX
XX 20-NOV-1997; 97WO-US021696.
XX

```

PR 20-NOV-1996; 96US-0031194P.
 PR 12-DEC-1996; 96US-003266P.
 PR 21-JUL-1997; 97US-0053200P.
 XX
 PA (YISS) YISSUM RES & DEV CO.
 PA (KOHN/) KOHN K I.
 XX
 PI Soreq H, Friedman A, Seldman S, Kauffer D;
 XX WPI; 1998-312172/27.
 DR
 XX
 PT Increasing the permeability of the blood/brain barrier - using e.g.
 PT adrenalectomy, atropine or acetylcholine esterase 14 splice variant peptide,
 PT useful for imaging and/or treatment of central nervous system disorders.
 XX
 PS Claim 1, 2; Page 45; 71pp; English.
 XX
 CC The present sequence represents human acetylcholine esterase-14 (AChE-14)
 CC readthrough splice variant. The protein sequence comprises residues
 CC encoded by exons 1-4 of human AChE followed by residues encoded by intron
 CC 4, while the last residue of the protein is encoded by exon 5 of AChE.
 CC The invention provides a pharmaceutical composition, for facilitating
 CC passage of compounds through the blood/brain barrier (BBB), comprising of
 CC AChE-14, 14 peptide or AChE-14 analogues together with a pharmaceutically
 CC acceptable carrier. The pharmaceutical composition is claimed to
 CC facilitate a reversible disruption of the BBB allowing transport of
 CC compounds through the BBB. The compounds, e.g. imaging agents,
 CC antibiotics or chemotherapeutic drugs, are claimed to be useful for the
 CC diagnosis and treatment of diseases or disorders of the CNS such as
 CC infections, neurochemical disorders, brain tumours, gliomas, etc
 CC
 SQ Sequence 600 AA;
 XX
 Query Match 100.0%; Score 147; DB 2; Length 600;
 Best Local Similarity 100.0%; Pred. No. 3; 1e-09; Indels 0; Gaps 0;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GMOGPAGSGWBERGSGSPGVTPLPSP 26
 Db 575 GMOGPAGSGWBERGSGSPGVTPLPSP 600
 XX
 RESULT 11
 ADCC3366
 ID ADCC3366 standard; protein; 145 AA.
 XX
 AC ADCC3366;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human novel config-encoded polypeptide sequence, SEQ ID NO:3448.
 XX
 KW Human, diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antihaemic; anticoagulant; thrombolytic; vulnary;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cyostatic;
 KW gene therapy.
 KW
 XX Homo sapiens.
 OS
 XX
 PN MO2003029271-A2.
 PD 10-APR-2003.
 XX
 PF 24-SEP-2002; 2002WO-US030474.
 XX
 PR 24-SEP-2001; 2001US-0324631P.
 XX
 PA (HYSE-) HYSEQ INC.
 XX

PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Dermanac RT;
 XX
 DR WPI; 2003-371981/35.
 XX
 DR N-PSDB; ADCC32599.
 XX
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 PS Example 2; SEQ ID NO 3448; 1185bp; English.
 XX
 CC The invention relates to 971 novel human cDNA sequences (ADCC29919-
 CC ADCC30889) and the polypeptides they encode (ADCC30890-ADCC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC cDNAs corresponding to the cDNA sequences of the invention
 CC (ADCC1861-ADCC32627) and the polypeptides encoded by the config (ADCC32628
 CC -ADCC3394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a human config-
 CC encoded polypeptide sequence used in an example of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIRO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 145 AA;
 XX
 Query Match 42.9%; Score 63; DB 7; Length 145;
 Best Local Similarity 54.2%; Pred. No. 8.3;
 Matches 13; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 3 GGPAGSGWBERGSGSPGVTPLPSP 26
 Db 71 GQDRGTGKGGSGSPGPTAGMMP 94
 XX
 RESULT 12
 ABB11475
 ID ABB11475 standard; peptide; 575 AA.
 XX
 AC ABB11475;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human R31449_3 homologue, SEQ ID NO:1845.
 XX
 KW Human, cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; actvlin;
 KW inhibin; chemokines; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;

chronic/inflammatory condition; proliferative retinopathy;
atherosclerosis; coronary heart disease; arterial ischemia;
bone disorder; osteoporosis; vascular growth disorder;
tissue regeneration; wound healing; infection; immune disorder;
cell culture; drug screening; gene therapy; anti-inflammatory;
antiaesthetic; antiarthritic; haemostatic; antitumescerotic;
cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
antifungal; vulnery; antitumor.

Homo sapiens.

WO200157188-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US003800.

03-FEB-2000; 2000US-00496914.
27-APR-2000; 2000US-00560875.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;
MPI; 2001-457740/49.
N-PSDB; ABA08719.

Human proteins and DNA encoding sequences useful for preventing, treating
or ameliorating a medical condition in a mammalian subject e.g. arthritis
and cancer.

Claim 20; Page 197; 1963pp; English.

Sequences ABB10961-ABE13330 represent 1350 novel human polypeptides, and
sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
invention also relates to vectors and recombinant host cells comprising a
nucleotide of the invention, methods of producing the novel polypeptides
or antibodies against the polypeptides, methods of detecting the nucleotides
or polypeptides in a sample, and methods of identifying compounds which
bind to polypeptides of the invention. Although novel, many of the
polypeptides of the invention have homology to known proteins, thereby
giving an insight into their probable biological activities, and hence
potential therapeutic applications. The polypeptides of the invention may
have various activities, including cytokine, cell proliferation or cell
differentiation activities; stem cell growth factor activity;
haematopoiesis regulatory activity; tissue growth activity;
immunomodulatory activity; activin- or inhibin-related activities;
chemotactic or chemokinetic activities; haemostatic, thrombotic or
thrombolytic activities; receptor or ligand activities; or may be
involved in oncogenesis, cancer cell proliferation or metastasis.
Depending on their biological activities, polypeptides and nucleotides of
the invention are useful for preventing, treating or ameliorating medical
conditions, e.g., by protein or gene therapy. Such conditions include
cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
proliferative retinopathy, atherosclerosis, coronary heart disease,
arterial ischemia, bone disorders (e.g., osteoporosis), and abnormal
vascular growth. Polypeptides involved with tissue regeneration and
repair (or nucleic acids encoding them) may be used to promote wound
healing (e.g., of burns, incisions and ulcers), while those with
immunomodulatory activities may be used in the treatment of viral,
bacterial and fungal infections in addition to immune disorders.
Polypeptides with growth factor activity may be used in cell cultures to
promote cell growth. For example, such polypeptides may be used to
manipulate stem cells in culture to give rise to neuroepithelial cells
that can be used to augment or replace cells damaged by illness,
autoimmune disease or accidental damage. The polypeptides and nucleotides
may also be used in the diagnosis of the above conditions, and in drug
screening techniques. The present sequence represents a novel human
polypeptide of the invention

Sequence 575 AA;

Query Match		Similarity	56.18%	Score	61.5	DB	4	Length	575;
Best Local		Similarity	56.18%	Pred.	No.	50;			
Matches	13;	Conservative	2;	Mismatches	7;	Indels	1;	Gaps	1
Oy	1	GMGQ-PAGSGWEGSGSPPTVP	22						
Dd	421	GGGGLPEGWGLEKSGELPPGI	443						
RESULT 13									
ID	AAB51865								
XX	AAB51865	standard; protein,	98 AA.						
AC	AAB51865;								
DT	16-FEB-2001	(first entry)							
XX									
DE	Human secreted protein sequence encoded by gene 39 SEQ ID NO:98.								
XX									
KM	Human; secreted protein; immunosuppressive; antiarthritic; antineumatic;								
KM	antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;								
KM	nootropic; neuroprotective; antibacterial; virucide; fungicide;								
KM	ophthalmological; vulnary; autoimmune disease; rheumatoid arthritis;								
KM	hyperproliferative disorders; cancer; cardiovascular disorder;								
KM	cardiac arrest; cerebrovascular disorder; nervous system disorder;								
KM	Alzheimer's disease; ocular disorder; wound healing; skin aging.								
OS	Homo sapiens.								
XX									
PN	WO200061626-A1.								
PD	19-OCT-2000.								
XX									
PF	06-APR-2000; 2000WO-US009066.								
XX									
PR	09-APR-1999; 99US-0128698P.								
XX	20-JAN-2000; 2000US-0176926P.								
XX									
PA	(HUMA-) HUMAN GENOME SCI INC.								
XX	(ROSE/) ROSEN C A.								
PI	Rosen CA, Ruben SM, Komatsoulis G;								
DR	WPJ; 2000-619227/59.								
XX	N-PSDB; AAC93517.								
PT	New nucleic acid molecules encoding 49 human secreted proteins for								
FT	diagnosing, preventing or ameliorating medical conditions and used for								
XX	food additives or preservatives.								
PS	Claim 11; Page 478; 516pp; English.								
XX									
CC	Polynucleotide sequences AAC93479 - AAC93527 represent cDNA encoding								
CC	human secreted proteins AAB51827 - AAB51875. Sequences AAB51876 -								
CC	AAB51927 represent alternative polypeptides encoded by the genes, and								
CC	amino acid sequences with which they share homology. The genes and								
CC	proteins have activities dependent on the tissues and cells in which they								
CC	are expressed. Examples of their activities include immunosuppressive;								
CC	antiarthritic; antineumatic; antiproliferative; cytostatic; cardiant;								
CC	vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;								
CC	virucide; fungicide; ophthalmological; and vulnary. The secreted								
CC	proteins, polynucleotides, antagonists and agonists may be useful in								
CC	treating, preventing and/or diagnosing diseases and disorders such as								
CC	autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative								
CC	disorders e.g. neoplasms of the breast or liver, cardiovascular disorders								
CC	e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,								
CC	angiogenesis, nervous system disorders e.g. Alzheimer's disease,								
CC	infections caused by bacteria, viruses and fungi and ocular disorders								
CC	e.g. cornal infection. The polypeptides can also be used to aid wound								
CC	healing and epithelial cell proliferation, to prevent skin aging due to								
CC	sunburn, to maintain organs before transplantation, for supporting cell								
CC	culture of primary tissues, to regenerate tissues and in chemotaxis. The								
CC	polypeptides can also be used as a food additive or preservative to								

Job time : 160 secs

2004

THIS PAGE IS BLANK

1	67	45.6	1378	2	097405	097405	haliotis di
2	67	45.6	1449	2	091000	091000	oncorhynch
3	62	42.2	234	2	06BP25	06BP25	debaromyce
4	62	42.2	640	2	06RQ4	06RQ4	mus muscul
5	62	42.2	640	2	BAD21403	BAD21403	mus muscul
6	62	42.2	699	2	06PIC4	06PIC4	mus muscul
7	62	42.2	659	2	Ah65148	Ah65148	mus muscul
8	62	42.2	1458	2	0910B9	0910B9	oncorhynch
9	60	40.8	334	1	FCN1_MOUSE	FCN1_MOUSE	mus muscul
10	60	40.8	1049	1	CA13_BOVIN	CA13_BOVIN	bos taurus
11	59.5	40.5	904	2	076271	076271	mytilus edi
12	59.5	40.5	905	2	08MW55	08MW55	mytilus gati
13	59	40.1	70	2	Q12985	Q12985	homo sapien
14	59	40.1	1447	2	06P4U1	06P4U1	brachydan
15	59	40.1	1447	2	06U153	06U153	brachydan
16	59	40.1	1447	2	Ah613249	Ah613249	brachydan
17	59	40.1	1447	2	AAR24536	AAR24536	brachydan
18	58.5	39.8	135	2	0951J6	0951J6	macaca fasc
19	58.5	39.8	191	2	09CSMO	09CSMO	mus muscul
20	58.5	39.8	386	2	Q011759	Q011759	pneumocyst
21	58.5	39.8	1076	2	Q001830	Q001830	pneumocyst
22	58	39.5	123	2	096ET3	096ET3	homo sapien
23	58	39.5	164	2	07XD78	07XD78	oryza sativ
24	58	39.5	185	2	07XD76	07XD76	oryza sativ
25	58	39.5	185	2	07XD77	07XD77	oryza sativ
26	58	39.5	185	2	0948R3	0948R3	oryza sativ
27	58	39.5	251	2	096MM4	096MM4	homo sapien
28	58	39.5	313	2	096P15	096P15	homo sapien
29	58	39.5	322	2	09N5K5	09N5K5	caenorhabd
30	58	39.5	322	2	0718J4	0718J4	homo sapien
31	58	39.5	333	2	AAQ88818	AAQ88818	homo sapien

32	58	39.5	421	2	09C0E3	09C0E3	homo sapiens
33	58	39.5	591	2	08C178	08C178	arabidopsis
34	58	39.5	703	1	CA28_HUMAN	P23067	homo sapiens
35	58	39.5	888	2	09O796	09O796	gallus galli
36	58	39.5	988	2	08CFM4	08CFM4	mus musculus
37	58	39.5	1222	2	08K173	08K173	mus musculus
38	58	39.5	1464	1	CA13_MOUSE	P08121	mus musculus
39	58	39.5	1464	2	07T732	07T732	mus musculus
40	58	39.5	1464	2	08BKX2	08BKX2	mus musculus
41	58	39.5	1464	2	08BLM4	08BLM4	mus musculus
42	57.5	39.1	196	2	Q25947	Q25947	plasmodium
43	57.5	39.1	280	2	08BP14	08BP14	mus musculus
44	57.5	39.1	280	2	08R138	08R138	mus musculus
45	57.5	39.1	2075	2	Q8TG44	Q8J644	fugu rubripes

ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT;	1378 AA.
097405			
ID	097405		
AC	097405;		
DT	01-MAY-1999 (TEMBRel. 10, Created)		
DT	01-MAY-1999 (TEMBRel. 10, last sequence update)		
DT	01-MAR-2004 (TEMBRel. 26, last annotation update)		
DE	Collagen pro alpha-chain precursor.		
GN	Name=Hdcol 1 alpha;		
OS	Halictus discus (abalone).		
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda		
OC	Veligastropoda; Halictoidae; Halictidae; Halictis.		
OX	NCBI_TaxID=36094;		

RP SEQUENCE FROM N.A.

RC TISSUE=Muscle;

RX MEDLINE=99234051; PubMed=10215888;

Yoneda C., Hirayama Y., Nakaya M., Matsubara Y., Irie S., Hatae K.,

RA Wacade S.!

RI "The occurrence of two types of collagen proalpha-chain in the abdominal

RI HALLOIS DISCUS MUSCLE.?
PI FIVE T DISCUS 361-314-731 (1999)

REL. BUL. U. BIOCHEM. 261:114-121 (1999).

DN	EMBL, F0017800, EMBL/3008:1, -:
DB	GO:000581, C:GO]agen: TRA

DB GO:0005737: C:cyton]asm: TRA
DA GO:0005581, C:cyton]asm: TRA:

DB: 60:0005201: F:ext rangel w/ ar mat

GO:0006817: P:phosphate transport

InterPro: IPR008160: Collagen.

InterPro: IPR000885: Fib collagen C

DR InterPro: IPR001007; VWF_C.

DR Pfam; PF01410; COLE1; 1.

DR Pfam; PF01391; Collagen; 17.

DR SMART; SM00038; COLFI; 1.

DR SMART; SM00214; VWC; 1.

DR PROSITE; PS50184; VWFC_2; 1.

KW Collagen; Signal.

FT	SIGNAL	1	18	Potent
----	--------	---	----	--------

SQ SEQUENCE 1378 AA; 132582 MW; 6E

Query Match	Score
45.6%	

Best Local Similarity 46.2%; Pred.

Matches 12; Conservative 3; Mises 1

1 234567891011121314151617181920212223242526272829303132333435363738394041424344454647484950515253545556575859606162636465666768697071727374757677787980818283848586878889909192939495969798991001011021031041051061071081091101111121131141151161171181191201211221231241251261271281291301311321331341351361371381391401411421431441451461471481491501511521531541551561571581591601611621631641651661671681691701711721731741751761771781791801811821831841851861871881891901911921931941951961971981992002012022032042052062072082092102112122132142152162172182192202212222232242252262272282292302312322332342352362372382392402412422432442452462472482492502512522532542552562572582592602612622632642652662672682692702712722732742752762772782792802812822832842852862872882892902912922932942952962972982993003013023033043053063073083093103113123133143153163173183193203213223233243253263273283293303313323333343353363373383393403413423433443453463473483493503513523533543553563573583593603613623633643653663673683693703713723733743753763773783793803813823833843853863873883893903913923933943953963973983994004014024034044054064074084094104114124134144154164174184194204214224234244254264274284294304314324334344354364374384394404414424434444454464474484494504514524534544554564574584594604614624634644654664674684694704714724734744754764774784794804814824834844854864874884894904914924934944954964974984995005015025035045055065075085095105115125135145155165175185195205215225235245255265275285295305315325335345355365375385395405415425435445455465475485495505515525535545555565575585595605615625635645655665675685695705715725735745755765775785795805815825835845855865875885895905915925935945955965975985996006016026036046056066076086096106116126136146156166176186196206216226236246256266276286296306316326336346356366376386396406416426436446456466476486496506516526536546556566576586596606616626636646656666676686696706716726736746756766776786796806816826836846856866876886896906916926936946956966976986997007017027037047057067077087097107117127137147157167177187197207217227237247257267277287297307317327337347357367377387397407417427437447457467477487497507517527537547557567577587597607617627637647657667677687697707717727737747757767777787797807817827837847857867877887897907917927937947957967977987998008018028038048058068078088098108118128138148158168178188198208218228238248258268278288298308318328338348358368378388398408418428438448458468478488498508518528538548558568578588598608618628638648658668678688698708718728738748758768778788798808818828838848858868878888898908918928938948958968978988999009019029039049059069079089099109119129139149159169179189199209219229239249259269279289299309319329339349359369379389399409419429439449459469479489499509519529539549559569579589599609619629639649659669679689699709719729739749759769779789799809819829839849859869879889899909919929939949959969979989991000100110021003100410051006100710081009101010111012101310141015101610171018101910201021102210231024102510261027102810291030103110321033103410351036103710381039104010411042104310441045104610471048104910501051105210531054105510561057105810591060106110621063106410651066106710681069107010711072107310741075107610771078107910801081108210831084108510861087108810891090109110921093109410951096109710981099110011011102110311041105110611071108110911101111111211131114111511161117111811191120112111221123112411251126112711281129113011311132113311341135113611371138113911401141114211431144114511461147114811491150115111521153115411551156115711581159116011611162116311641165116611671168116911701171117211731174117511761177117811791180118111821183118411851186118711881189119011911192119311941195119611971198119912001201120212031204120512061207120812091210121112121213121412151216121712181219122012211222122312241225122612271228122912301231123212331234123512361237123812391240124112421243124412451246124712481249125012511252125312541255125612571258125912601261126212631264126512661267126812691270127112721273127412751276127712781279128012811282128312841285128612871288128912901291129212931294129512961297129812991300

1 GMLGGPAGSGWEEGSGSPFGVLPFST

nb
1111 CTTCBACBCBDBCBDBGVGBVSE

DD YYY GHIJFGFDFGFFGFDFGFIJGFVIG

RESULT 2

0910C0

ID 0910C0 PRELIMINARY: PRT:

AC Q910C0;

DT 01-DEC-2001 (TrEMBLrel. 19, Created

DT 01-DEC-2001 (Tremblay, 19, Last seen)

```

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Collagen a1(I).
GN Name=COL1A1.
OS Oncochrychus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei;
OC Procacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
NCBI_Taxid=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21257802; PubMed=11358457;
RA Saito M., Takenouchi Y., Kunitaki N., Kimura S.,
RT "Complete primary structure of rainbow trout type I collagen
RT consisting of a1(I)2(I)3(I) heterotrimers.",
RT Eur. J. Biochem. 268:2817-2827(2001).
DR EMBL: AB052835; BAB5561.1;
DR GO: GO:0005581; C:collagen; IEA.
DR GO: GO:0005737; C:cytoplasm; IEA.
DR GO: GO:0005201; F:extracellular matrix structural constituent; IEA.
DR GO: GO:0006817; P:phosphate transport; IEA.
DR InterPro: IPR008161; Clg_helix.
DR InterPro: IPR008160; Collagen.
DR InterPro: IPR00885; F1b_collagen_C.
DR Pfam: PF01410; COLFI; 1.
DR Pfam: PF01391; Collagen; 18.
DR ProDom: PD000007; Clg_helix; 4.
DR ProDom: PD002078; F1b_collagen_C; 1.
DR SMART: SM00038; COLFI; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01208; VWC_1; 1.
DR PROSITE: PS50184; VWC_2; 1.
KW Collagen.
SQ SEQUENCE 1449 AA; 137116 MW; 62BEBFA7BFD55288 CRC64;

Query Match 45.6%; Score 67; DB 2; Length 1449;
Best Local Similarity 46.2%; Pred. No. 31;
Matches 12; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 GMOGPGSGWEGSGSPPGVTPFSP 26
DB 608 GVAAGPGERGEGQAGGPGFQGLSGP 633

RESULT 3
Q6BP25 PRELIMINARY; PRT; 234 AA.
AC Q6BP25;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similarity.
GN ORFNames=DEHA0517853g;
OC Debaryomyces hanseni (Yeast) (Torulaspora hanseni1).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
NCBI_Taxid=4959;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CHS767;
RG GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Badoir A., Barde V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Bolaram L., Boyer J., Catolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Gropi A.,
RA Hantreave F., Henneguain C., Jaumaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Niclaud J.M., Nikolski M., Ozdas S., Ozler-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,

```

```

RA Zenlou-Meyer M., Zivanovic I., Bojotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudon B., Scarpelli C., Galliard C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.",
RT Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CHS767;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL: CR82137; CG88303.1;
SQ SEQUENCE 234 AA; 22636 MW; 3C0B851B5D0095C9 CRC64;

Query Match 42.2%; Score 62; DB 2; Length 234;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GMOGPGSGWEGSGSPPGV 20
DB 126 GQGPKGAGGERGIGGPGV 145

RESULT 4
Q6KAQ4 PRELIMINARY; PRT; 640 AA.
AC Q6KAQ4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLU00201 protein (Fragment).
GN Name=mFLU00201;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=embryonic tail;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hirooka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of 110 Mouse Homologous cDNAs
RT Identified by Screening of Terminal Sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries.",
RL DNA Res. 11:167-180(2004).
DR EMBL: AK31153; BAD21403.1;
DR InterPro: IPR001073; Clg.
DR InterPro: IPR008160; Collagen.
DR InterPro: IPR008983; TNF_like.
DR Pfam: PF00386; Clg; 1.
DR Pfam: PF01391; Collagen; 7.
DR PRINTS: PR00007; COMPLEMENTC1Q.
DR SMART: SM00110; Clq; 1.
DR PROSITE: PS01113; Clq; 1.
KW Collagen.
FT NON TER.
SQ SEQUENCE 640 AA; 61034 MW; 75CC9DBA85AC4B5 CRC64;

Query Match 42.2%; Score 62; DB 2; Length 640;
Best Local Similarity 60.0%; Pred. No. 52;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GMOGPGSGWEGSGSPGV 20
DB 307 GRRGPGSGKEVGPQGPV 326

RESULT 5
BAD21403 PRELIMINARY; PRT; 640 AA.
AC BAD21403;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)

```

DT 01-JUN-2004 (TrEMBLrel. 27, last annotation update)
 DE MFLJ00201 protein (Fragment).
 GN MFLJ00201.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBITaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic tail;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
 RT Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.,
 RT "Prediction of the Coding Sequences of Mouse Homologues of Flj Gene:
 RT The Complete Nucleotide Sequences of 110 Mouse Flj-homologous cDNAs
 RT Identified by Screening of Terminal sequences of cDNA Clones Randomly
 RT Sampled from Size-fractionated Libraries."
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK131153; BAD21403.1; -.
 FT NON TER 1 1
 SQ SEQUENCE 640 AA; 61034 MW; 75CC9DEBA5AC4B5 CRC64;

 Query Match 42.2%; Score 62; DB 2; Length 640;
 Best Local Similarity 60.0%; Pred. No. 52;
 Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

 QY 1 GMOGPAGSGMEBSSGSPGV 20
 DB 307 GRGPPGSKGEVGP GPGPV 326

 RESULT 6
 ID 06P1C4 PRELIMINARY; PRT; 699 AA.
 AC 06P1C4;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Procollagen, type VIII, alpha 2.
 GN Name=Col8a2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBITaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RA Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC065148; AAH65148.1; -.
 DR

DR InterPro: IPR001073; Clq.
 DR InterPro: IPR008150; Collagen.
 DR InterPro: IPR008983; TNF-like.
 DR Pfam: PF00386; Clq; 1.
 DR Pfam: PF01391; Collagen; 7.
 DR PRINTS: PRO0007; COMPLEMENTC1Q.
 DR SMART: SM00110; Clq; 1.
 DR PROSITE: PS01113; Clq; 1.
 KW Collagen.
 SQ SEQUENCE 699 AA; 66943 MW; FCBD4FBB44642646 CRC64;

 Query Match 42.2%; Score 62; DB 2; Length 699;
 Best Local Similarity 60.0%; Pred. No. 57;
 Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

 QY 1 GMOGPAGSGMEBSSGSPGV 20
 DB 366 GRGPPGSKGEVGP GPGPV 385

 RESULT 7
 ID AAH65148 PRELIMINARY; PRT; 699 AA.
 AC AAH65148;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DE 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Procollagen, type VIII, alpha 2.
 GN COL8A2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBITaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RA Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC065148; AAH65148.1; -.
 DR

```

RESULT 8
Q910B9 PRELIMINARY; PRT; 1458 AA.
ID Q910B9;
AC Q910B9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Collagen a3 (I).
GN Name=COL1A3;
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Procarangopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OC NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21257802; PubMed=11358497;
RX Saito M., Takenouchi Y., Kunitaki N., Kimura S.;
RT "Complete primary structure of rainbow trout type I collagen
RL Eur. J. Biochem. 268:2817-2827(2001).
DR EMBL; AB052836; BAB55662.1; -
DR GO; GO:0005581; C:collagen; IEA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; C1g_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 18.
DR ProDom; PD000007; C1g_helix; 2.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00338; COLFI; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01208; VWF_C_1; 1.
DR PROSITE; PS50184; VWF_C_2; 1.
DR KEGG; COLLAGEN.
SQ SEQUENCE 1458 AA; 137757 MW; AB1F9F3410A98650 CRC64;

Query Match 42.2%; Score 62; DB 2; Length 1458;
Best Local Similarity 63.2%; Pred. No. 1.2e+02;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 GNGPAGSGMEGSGSPG 19
Db 611 GNGPAGSGMEGSGSPG 629

RESULT 9
FCN1_MOUSE STANDARD; PRT; 334 AA.
ID FCN1_MOUSE;
AC 070165;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ficolin 1 precursor (Collagen/fibrinogen domain-containing protein 1)
DE (Ficolin-A) (Ficolin A) (M-Ficolin).
GN Name=FCN1; Synonyms=Fena;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C; Tissue=Liver;
RX MEDLINE=98205801; PubMed=9535745;
RX Fujimori Y., Harumiya S., Fukumoto Y., Miura Y., Yagasaki K.,
RX Tachikawa H., Fujimoto D.;
RT "Molecular cloning and characterization of mouse ficolin-A.";

```

```

RL Biochem. Biophys. Res. Commun. 244:796-800(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RX Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RX Altschul S.F., Zeeberg B., Burow K.H., Scheffer C.F., Bat N.K.,
RX Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RX Datchenko L., Marusheva K., Farmer A.A., Rubin G.M., Hong L.,
RX Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RX Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RX Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RX Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RX Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huily S.W.,
RX Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RX Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RX Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RX Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RX Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RX Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RX Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Involved in serum exerting lectin activity. Binds GlcNAc
CC (By similarity).
CC -1- SUBUNIT: Homopolymer. Interacts with elastin (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Found on the monocyte surface (By
CC similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in liver and spleen.
CC -1- SIMILARITY: Belongs to the ficolin lectin family.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB007813; BAA25126.1; -
DR EMBL; BC019180; AAH19180.1; -
DR PIR; JCS980; JCS980.
DR HSP; P02671; 1FZD.
DR WGI; WGI:1340905; Fcna.
DR InterPro; IPR008161; C1g_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF01391; Collagen; 1.
DR Pfam; PF00147; Fibrinogen_C; 1.
DR ProDom; PD000007; C1g_helix; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
DR Collagen; Glycoprotein; Lectin; Multigene family; Repeat; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 334 Ficolin 1.
FT DOMAIN 50 88 Collagen-like.
FT DOMAIN 152 298 Fibrinogen C-terminal.
FT CARBOHYD 261 261 N-linked (GlcNAc...) (potential).
SQ SEQUENCE 334 AA; 36298 MW; 9D30C05036AA04B1 CRC64;

Query Match 40.8%; Score 60; DB 1; Length 334;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GNGPAGSGMEGSGSPGV 20
Db 77 GNGPAGSGMEGSGSPGV 96

RESULT 10

```

ID	Ca13_BOVIN	STANDARD	PRT	1049 AA.
AC	P04258			
DT	20-MAR-1987	(Rel. 04, Created)		
DT	20-MAR-1987	(Rel. 04, Last sequence update)		
DE	05-JUL-2004	(Rel. 44, Last annotation update)		
DE	Collagen alpha 1(III) chain.			
CN	Name-COL1A1			
OS	Bos taurus (Bovine)			
OC	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovinae; Bos.			
OX	NCBI_TaxId=9913			
RN	[1]			
RP	SEQUENCE OF 1-242.			
RX	MEDLINE=80026026; PubMed=488906;			
RA	Fietzek P.P., Allmann H., Rautenberg J., Henkel W., Wachter E.,			
RA	Kuhn K.;			
RT	"The covalent structure of calf skin type III collagen. I. The amino			
RT	acid sequence of the amino terminal region of the alpha 1(III) chain			
RL	(positions 1-222)."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 360:809-820(1979).			
RN	[2]			
RP	SEQUENCE OF 243-422.			
RX	MEDLINE=80026027; PubMed=488907;			
RA	Dewes H., Fietzek P.P., Kuhn K.;			
RT	"The covalent structure of calf skin type III collagen. II. The amino			
RT	acid sequence of the cyanogen bromide peptide alpha 1(III)CB1,8,10,2			
RL	(positions 223-402)."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 360:821-832(1979).			
RN	[3]			
RP	SEQUENCE OF 423-571.			
RX	MEDLINE=80026028; PubMed=488908;			
RA	Bentz H., Fietzek P.P., Kuhn K.;			
RT	"The covalent structure of calf skin type III collagen. III. The amino			
RT	acid sequence of the cyanogen bromide peptide alpha 1(III)CB4			
RL	(positions 403-551)."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 360:833-840(1979).			
RN	[4]			
RP	SEQUENCE OF 572-809.			
RX	MEDLINE=80026029; PubMed=488909;			
RA	Lang H., Glanville R.W., Fietzek P.P., Kuhn K.;			
RT	"The covalent structure of calf skin type III collagen. IV. The amino			
RT	acid sequence of the cyanogen bromide peptide alpha 1(III)CB5			
RL	(positions 552-788)."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 360:841-850(1979).			
RN	[5]			
RP	SEQUENCE OF 809-947.			
RX	MEDLINE=80026030; PubMed=488910;			
RA	Dewes H., Fietzek P.P., Kuhn K.;			
RT	"The covalent structure of calf skin type III collagen. V. The amino			
RT	acid sequence of the cyanogen bromide peptide alpha 1(III)CB9A			
RL	(position 789-927)."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 360:851-860(1979).			
RN	[6]			
RP	SEQUENCE OF 948-1049.			
RX	MEDLINE=80026031; PubMed=488911;			
RA	Allmann H., Fietzek P.P., Glanville R.W., Kuhn K.;			
RT	"The covalent structure of calf skin type III collagen. VI. The amino			
RT	acid sequence of the carboxyterminal cyanogen bromide peptide alpha			
RL	1(III)CB9B (positions 928-1028)."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 360:861-868(1979).			
CC	-1- FUNCTION: Collagen type III occurs in most soft connective tissues			
CC	along with type I collagen.			
CC	-1- SUBUNIT: Trimers of identical alpha 1(III) chains. The chains are			
CC	linked to each other by interchain disulfide bonds. Trimers are			
CC	also cross-linked via hydroxylysines.			
CC	-1- PM: Prolines at the third position of the tripeptide repeating			
CC	unit (G-X-Y) are hydroxylated in some or all of the chains.			
CC	-1- SIMILARITY: Belongs to the fibrillar collagen family.			
DR	PIR: A02862; CGB07S.			
DR	InterPro: IPR008161; Clg_helix.			
DR	InterPro: IPR008160; Collagen.			

DR	InterPro: IPR001007; VWF_C.
DR	Pfam: PF01391; Collagen_17.
DR	ProDom: PD000007; C1g_helix; 3.
DR	PROSITE, PS01208; VWF_C_1; PARTIAL.
KM	Collagen; Connective tissue; Direct protein sequencing;
KM	Extracellular matrix; Glycoprotein; Hydroxylation; Repeat.
FT	DOMAIN 1 14 Nonhelical region (N-terminal).
FT	DOMAIN 15 1040 Triple-helical region.
FT	DOMAIN 1041 1049 Nonhelical region (C-terminal).
FT	MOD_RES 95 95 5-hydroxylysine.
FT	MOD_RES 107 107 5-hydroxylysine.
FT	MOD_RES 119 119 5-hydroxylysine.
FT	MOD_RES 938 938 5-hydroxylysine.
FT	MOD_RES 950 950 5-hydroxylysine.
FT	CARBOHYD 107 107 O-linked (Gal. . .).
FT	CARBOHYD 950 950 O-linked (Gal. . .).
FT	DISULFID 1040 1040 Interchain.
FT	DISULFID 1041 1041 Interchain.
SEQ	SEQUENCE 1049 AA; 93651 MW; 8BEC33D1C66EC9A3 CRC64;
Query Match	
Best Local Similarity	40.8%; Score 60; DB 1; Length 1049;
Matches 12; Conservative	57.1%; Pred. No. 1.5e+02;
	0; Mismatches 9; Indels 0; Gaps 0;
Qy	1 GMCGPAGSGWEEGSGSPPGVT 21
Db	447 GPCGPAGKNGERTGPCGPGPT 467
RESULT 11	
ID 076271	PRELIMINARY; PRT; 904 AA.
AC 076271;	
DT 01-NOV-1998 (TREMBLrel. 08; Created)	
DT 01-NOV-1998 (TREMBLrel. 08; Last sequence update)	
DT 01-OCT-2003 (TREMBLrel. 25; Last annotation update)	
DE Nongradient byssal.	
OS Mytilus edulis (Blue mussel).	
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;	
OC Mytiloidea; Mytilidae; Mytilus.	
OX NCBI_TaxID=6550;	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=98393676; PubMed=9724735;	
RA Qin X.X., Waite J.H.;	
RT "A potential mediator of collagenous block copolymer gradients in	
RT mussel byssal threads.";	
RL Proc. Natl. Acad. Sci. U.S.A. 95:10517-10522 (1998).	
DR EMBL; AF043944; AAC33847.1; -.	
DR GO; GO:0005737; C:cytoplasm; IEA.	
DR GO; GO:0006817; P:phosphate transport; IEA.	
DR InterPro: IPR008161; C1g_helix.	
DR InterPro: IPR008160; Collagen.	
DR Pfam; PF01391; Collagen; 7.	
DR ProDom; PD000007; C1g_helix; 1.	
KM Collagen.	
SEQ SEQUENCE 904 AA; 77883 MW; 5529135651AD4C40 CRC64;	
Query Match	
Best Local Similarity	40.5%; Score 59.5; DB 2; Length 904;
Matches 13; Conservative	50.0%; Pred. No. 1.5e+02;
	3; Mismatches 9; Indels 1; Gaps 1;
Qy	1 GMCGPAGSGWEEGSGSPPGVTPLSP 26
Db	227 GPRGPAGPPDQHGSGPPG-PPGNSP 251
RESULT 12	
ID 08MWS5	PRELIMINARY; PRT; 905 AA.
AC 08MWS5;	
DT 01-OCT-2002 (TREMBLrel. 22; Created)	
DT 01-OCT-2002 (TREMBLrel. 22; Last sequence update)	

```

DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)
DE Precollagen-NG.
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxId=29158;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22038007; PubMed=12042339;
RA Lucas J.M., Vaccaro E., Waite J.H.;
RT "A molecular, morphometric and mechanical comparison of the structural
RT elements of byssus from Mytilus edulis and Mytilus
RT galloprovincialis."
RL J. Exp. Biol. 205:1807-1817(2002).
DR EMBL; AF48524; AAM34599.1; -
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 7.
KW Collagen.
SQ SEQUENCE 905 AA; 79251 MW; 499F6FB26A5C182 CRC64;

Query Match 40.5%; Score 59.5; DB 2; Length 905;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 13; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 1 GMOGPAGSGMBSGSPGVTPLFSP 26
DB 232 GPGPPGAPGEGKRGARGPVGIVGPIGPP 67

RESULT 13
ID 012985 PRELIMINARY; PRT; 70 AA.
AC 012985;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, last annotation update)
DE Alpha-1 type II collagen (Fragment).
GN Name=COL2A1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Blood;
RX MEDLINE=95150028; PubMed=7847372;
RA Tiller G.E., Weis M.A., Polunbo P.A., Gruber H.E., Rimoin D.L.,
RA Cohn D.H., Eyre D.R.;
RT "An RNA-splicing mutation (G>51VS20) in the type II collagen gene
RT (COL2A1) in a family with spondyloepiphyseal dysplasia congenita.";
RL Am. J. Hum. Genet. 56:388-395 (1995).
DR EMBL; U15195; AAB60370.1; -.
DR PIR; A38513; CGHUC.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 1.
DR Prodom; PD000007; C1g_helix; 1.
KW Collagen.
FT NON TER
FT NON TER
SQ SEQUENCE 70 AA; 6482 MW; 13AE3044C3F2FC07 CRC64;

Query Match 40.1%; Score 59; DB 2; Length 70;
Best Local Similarity 46.7%; Pred. No. 12;
Matches 14; Conservative 1; Mismatches 11; Indels 4; Gaps 1;

QY 1 GMOGPAGSGMBSGSPGVTPLFSP 26
DB 38 GPGPPGAPGEGKRGARGPVGIVGPIGPP 67

```

```

RESULT 14
ID 06P4U1 PRELIMINARY; PRT; 1447 AA.
AC 06P4U1;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, last annotation update)
DE Collal protein.
GN Name=collal;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Cartinci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., Mesman P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalski U., Smalios D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Embryo;
RX Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC063248; AAB63249.1; -.
DR InterPro; IPR008161; C1g_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002181; Fibrinogen_C.
DR InterPro; IPR000885; Fib_collagen_C.
DR InterPro; IPR009041; FMP_SGCI.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 18.
DR Pfam; PF00093; VWC; 1.
DR Prodom; PD000007; C1g_helix; 1.
DR Prodom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01208; VWF_C_1; 1.
DR PROSITE; PS50184; VWF_C_2; 1.
KW Collagen.
SQ SEQUENCE 1447 AA; 136955 MW; 74723FCBACAAD86 CRC64;

Query Match 40.1%; Score 59; DB 2; Length 1447;
Best Local Similarity 46.2%; Pred. No. 2.7e+02;
Matches 12; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 GMOGPAGSGMBSGSPGVTPLFSP 26
DB 607 GPGPPGAPGEGKRGARGPVGIVGPIGPP 632

```


RESULT 15

```

ID 06UJ5 PRELIMINARY; PRT; 1447 AA.
AC 06UJ5
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Chihuahua.
GN Name-chi;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22984997; PubMed=14623232;
RA Fisher S., Jagadeeswaran P., Halpern M.E.;
RT "Radiographic analysis of zebrafish skeletal defects.";
RL Dev. Biol. 264:64-76(2003).
DR EMBL; AY380817; AAR24536.1; -.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002181; Fibrinogen_C.
DR InterPro; IPR000885; Fib collagen_C.
DR InterPro; IPR009041; BMP_SGCT.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 17.
DR Pfam; PF00093; VWC; 1.
DR ProDom; PD002078; Fib collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01208; VWF_C_1; 1.
DR PROSITE; PS50184; VWF_C_2; 1.
DR KW Collagen.
SQ SEQUENCE 1447 AA; 137144 MW; 9CABD561F5BA36BF CRC64;

```

```

Query Match 40.1%; Score 59; DB 2; Length 1447;
Best Local Similarity 46.2%; Pred. No. 2.7e+02;
Matches 12; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

```

```

QY 1 GMOGPAGSGWBEQSGSPGVTPPLFSP 26
DB 607 GPAGPAGRGEGGAGACPFQGLPGP 632

```

Search completed: December 29, 2004, 15:02:08
 Job time : 201 secs

THIS PAGE IS BLANK